Psallidas, I., Kanellakis, N. I., Gerry, S., Thézénas, M. L., Charles, P. D., Samsonova, A., Schiller, H. B., Fischer, R., Asciak, R., Hallifax, R. J., Mercer, R., Dobson, M., Dong, T., Pavord, I. D., Collins, G. S., Kessler, B. M., Pass, H. I., Maskell, N., Stathopoulos, G. T., & Rahman, N. M. (2018). Development and validation of response markers to predict survival and pleurodesis success in patients with malignant pleural effusion (PROMISE): a multicohort analysis. *Lancet Oncology*, *19*(7), 930-939. https://doi.org/10.1016/S1470-2045(18)30294-8

Peer reviewed version

Link to published version (if available):
10.1016/S1470-2045(18)30294-8

Survival and pleurodesis response markers in malignant pleural effusion - The PROMISE study: a multicohort analysis

Ioannis Psallidas, PhD[1,2,3], Nikolaos I. Kanellakis, MSc[1,2,3], Stephen Gerry, PhD[4], Marie Laëtitia Thézénas, MSc[5], Philip D. Charles, PhD[5], Anastasia Samsonova, PhD[6,7], Herbert B. Schiller, PhD[8], Roman Fischer, PhD[5], Rachelle Asciak, MD[1,2,3], Robert J. Hallifax, MRCP[1,2,3], Rachel Mercer, MRCP[1,2,3], Melissa Dobson, BSc[3], Prof Tao Dong, PhD[9,10], Prof Ian D. Pavord, PhD[1], Prof Gary S. Collins, PhD[4], Prof Benedikt M Kessler, PhD[5], Prof Harvey I. Pass, PhD[11], Prof Nick Maskell, DM[12], Georgios T. Stathopoulos, PhD[8,13], Najib M Rahman, DPhil[1,2,3,14]

Affiliations:

1. Oxford Centre for Respiratory Medicine, Churchill Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, OX3 7LE, UK
2. Laboratory of Pleural and Lung Cancer Translational Research, Nuffield Department of Medicine, University of Oxford, Oxford, OX3 7FZ, UK
3. Oxford Respiratory Trials Unit, Nuffield Department of Medicine, University of Oxford, Oxford, OX3 7LE, UK
4. Centre for Statistics in Medicine, Botnar Research Centre, University of Oxford, Oxford, OX3 7LD, UK
5. Mass Spectrometry Laboratory, Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, OX3 7FZ, UK
6. Bioinformatics Core, Department of Oncology, University of Oxford, Oxford, OX3 7DQ, UK
7. Institute of Translational Biomedicine, St. Petersburg State University, St Petersburg 199034, Russia
8. Comprehensive Pneumology Center (CPC) and Institute for Lung Biology and Disease (iLBD), University Hospital, Ludwig-Maximilian University (LMU) and Helmholtz Center Munich, Member of the German Center for Lung Research (DZL); Max-Lebsche-Platz 31, 81377, Munich, Bavaria, Germany

9. Centre for Translational Immunology, CAMS-Oxford Institute, Nuffield Department of Medicine Research Building, University of Oxford, Oxford OX3 9FZ, UK
10. MRC Human Immunology Unit, Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DS
11. Department of Cardiothoracic Surgery New York University Langone Medical Center, New York, NY 10016 USA
12. Academic Respiratory Unit, School of Clinical Sciences, University of Bristol, Bristol, BS10 5NB, UK
13. Laboratory for Molecular Respiratory Carcinogenesis, Department of Physiology, Faculty of Medicine; University of Patras, 1 Asklepiou Str., Patras, 26504 Greece.
14. National Institute for Health Research Oxford Biomedical Research Centre, University of Oxford, Oxford, OX3 7LE, UK

Corresponding author:

Ioannis Psallidas, PhD

Oxford University NHS Foundation Trust, Old Road, Churchill site, OX3 7LE, Oxford, United Kingdom

Email: ioannis.psallidas@ndm.ox.ac.uk

Telephone: +44 (0) 1865257104 Fax: +44(0) 1865 857109

SUMMARY

Background: Malignant pleural effusion (MPE) is increasing worldwide and prognostic biomarkers to plan treatment and to understand the underlying mechanisms of disease progression remain unidentified. The PROMISE study was designed with the objectives to discover, validate, and prospectively assess biomarkers of survival and pleurodesis response in MPE and build a score that forecasts survival.

Methods: Five separate and independent datasets from randomized controlled trials (*n* = 597 patients) were used to investigate potential biomarkers of survival and pleurodesis. Mass spectrometry based discovery was used to investigate pleural fluid samples for differential protein expression in a discovery patient groups with different survival and pleurodesis outcomes. Clinical, radiological and biological variables were entered into least absolute shrinkage and selection operator regression to build a model that predicts 3-month mortality. The model was evaluated using internal and external validation.

Findings: 16 and 7 biomarker candidates of survival and pleurodesis respectively were identified in the discovery dataset. Three independent datasets (*n* = 435) were used for biomarker validation. All pleurodesis biomarkers failed while GSN, MIF, VCAN and TIMP1 emerged as accurate predictors of survival. Eight variables (haemoglobin, c-reactive protein, white blood cell count, performance status, cancer type, pleural fluid TIMP1 levels and previous chemoradiotherapy) were validated and used to develop a survival score. Internal validation using bootstrap resampling and external validation using 162 patients from two independent datasets demonstrated good discrimination (c-index 0·78 and 0·89 respectively).

Interpretation: The PROMISE score is the first prospectively validated prognostic model for MPE that combines biological and clinical parameters to accurately estimate 3-month mortality.

Funding: European Respiratory Society, Medical Research Funding, Slater & Gordon Research Fund and Oxfordshire Health Services Research Committee Research Grants.

Key words: Malignant pleural effusion; metastasis; biomarker; mortality; cancer

Word count abstract: 276

RESEARCH IN CONTEXT

Evidence before this study:

We searched PubMed for systematic reviews and studies with high throughput screening, retrospective analysis, development of clinical score and prospective validation of survival and treatment biomarkers in malignant pleural effusion published up to 5th of December, 2017, with the keywords “biomarkers”, “malignant pleural effusion (MPE)” ‘’survival’’ and ‘’pleurodesis’’. We applied search filters for “adult” and “English”. There were no studies that have been published and met our criteria.

Currently the factors that drive and define malignant progression, resistance to therapy, and mortality in MPE are poorly understood. A prognostic model using clinical information in combination with blood and/or pleural fluid biomarkers that predicts survival and pleurodesis response would be of importance for clinical management and improve our understanding of MPE pathobiology

Added value of this study

This study, to our knowledge, is the first translational assessment combining discovery technology with high-quality clinical data from five different datasets for the development of a prognostic score for MPE. The results demonstrate the discovery, assessment and prospective validation of a novel prognostic score, including clinical data with pleural fluid biomarkers.

Implications of all the available evidence

There is now a compelling need for stratification of management of patients with MPE. A prognostic score predicting the 3-month risk of death will provide important information about patient prognosis and guide the selection of appropriate

management strategies. Additionally our results identify four biological pathways with the potential to guide novel treatments of MPE.

INTRODUCTION

Malignant pleural effusion (MPE) is a common condition that annually affects 500-700 individuals per million population[1]. The global cancer rate is rising, and with improvements in systemic therapy that allow many patients to live longer, the burden of MPE is increasing[2-4]. In addition to the clinical significance of MPE, dissemination of cancer cells in the pleural cavity is a biologic hallmark of highly metastatic malignancy regardless of the primary neoplasm, and current guidelines quote a median survival of three to 12 months[1]. Pleurodesis, the most common procedure used to treat MPE, prevents fluid accumulation through induction of pleural inflammation and fibrosis, succeeding in approximately 70% of patients assessed at one month[1,5].

In recent years, cancer treatment and pleural procedures available for MPE have expanded, leading to improved stratification of patients and progressive individualization of treatment, but also to diversification of outcomes such as pleurodesis success and survival[6]. However, the factors that drive and define malignant progression, resistance to therapy, and mortality in MPE are poorly understood. To this end, a prognostic model using clinical information in combination with blood and/or pleural fluid biomarkers that predicts survival and pleurodesis response would be of importance for clinical management, and would improve our understanding of MPE pathobiology. Such a risk stratification system was recently proposed: the LENT score predicts patients' survival based on tumour type, pleural fluid lactate dehydrogenase (LDH), Eastern Cooperative Oncology Group (ECOG) performance score, and blood neutrophil/lymphocyte ratio[7]. The LENT was

developed using clinical data from three existing patient cohorts combined with biochemical measurements of pleural fluid and has not been prospectively assessed.

This study reports the development and validation of the survival and pleurodesis response markers in MPE (PROMISE), the first score dedicated to the prospective assessment of the survival of patients with MPE simultaneously employing clinical, biological, and radiological parameters. Data and samples from patients recruited to previously published multinational, multicentre clinical trials with identical and established recruitment criteria were used.

MATERIALS AND METHODS

The study consisted of a three-step approach (discovery, validation, prospective assessment) that included Label Free Quantitative proteomic analyses of cell-free pleural fluid from a prospectively recruited cohort (biologic discovery), identification of prognostic clinical variables in published cohorts (clinical discovery), model development, model validation in two separate cohorts, and finally prospective assessment in two independent datasets from ongoing studies randomizing patients with MPE.

Discovery and internal validation databases

Clinical data and pleural fluid samples were used from three previous multicentre clinical trials, TIME-1[8], TIME-2[9], and TIME-3[10], which were prospectively collected between 2007 and 2016. Patients at the first onset of MPE were recruited on identical, clinically robust criteria from hospitals in the UK, USA, and Canada. Consistent data collected for each study participant included type of primary cancer, pleurodesis outcome, chest X-ray findings, pleural fluid LDH, performance status (ECOG), full blood count, haemoglobin, previous treatment (either with first line chemotherapy and/or radiotherapy), forced expiratory volume 1 (FEV_1), and C-reactive protein (CRP). Survival data for each patient was collected from the trials' databases and cross-validated with a national registry.

External validation databases

Individuals from the Oxford Radcliffe Pleural Biobank and SIMPLE study[11] were used for external validation of the prognostic score. Pleural fluid samples were prospectively collected alongside clinical data. In total, 158 samples were used for

the validation. The exact survival data for each patient was collected from the trial databases and cross-validated with a national registry.

Pleural fluid analysis and mass spectrometry

For biomarker discovery, a label free quantitative proteomic analysis (immunodepletion followed by Gel-Aided Sample Preparation (GASP)[12] and liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed on TIME-2 samples (total 114 patients) from 39 patients for the survival analysis (20 with poor and 19 with good survival) and from 26 patients (11 with failure and 15 with pleurodesis success) for the pleurodesis analysis. Samples were processed by depleting the top 12 high abundant proteins prior to digestion and analysed as per protocol (supplementary appendix page 3). Peptides were analysed by LC-MS/MS using a Q Exactive mass spectrometer (Thermo Scientific, Bremen, Germany) coupled with ultra-high performance liquid chromatography (Dionex Ultimate 3000 UHPLC, Thermo) using a PepMap RSLC column (C18, 2 um, 100 A, 75 um x 50 cm).

Data were quantified in Progenesis QI for Proteomics (Nonlinear Dynamics). Peptide-spectrum matching was performed using Mascot (Matrix Science, UK) against the UniProt SwissProt database (retrieved 26/11/15) restricted to Homo sapiens taxonomy. Searches were performed using the following universal parameters - variable modifications: oxidation (M), propionamide (K), propionamide (N-term), fixed modifications: propionamide (C), precursor tolerance: 10 ppm, fragment tolerance: 0·05 Da, maximum missed cleavages: 1, up to 1 C13 Peaks considered. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset

identifier PXD008682. For further information please refer to the online supplementary appendix (page 3).

Bioinformatics analysis of proteomic data

Peptides were assigned to a specific protein, and different isoforms matched to the same protein. Three different bioinformatics groups analysed the data independently. R software (version 3) was used for the analysis of LC-MS/MS data[13]. The Bioconductor (release 3·6) packages Biobase[14], limma[15], and qvalue[16] were used. Proteins identified with only one unique peptide and proteins detected in only three samples per condition were excluded from further analysis. Protein counts were log2 transformed and subsequently normalised using the cyclicLoess algorithm. Independent *t* tests were used to identify statistically significant signals.

Ingenuity pathway analysis

To discover biological pathways related to survival and pleurodesis outcome the statistically significant proteins was subjected to pathway and upstream regulator analysis using the Ingenuity Pathway Analysis (IPA, spring release 2016) software (Qiagen Bioinformatics).

Mass spectrometry protein validation with Elisa and Luminex

Elisa and Luminex protein assays were performed to measure pleural fluid protein levels. Each protein was measured in technical triplicates per sample

Statistical methodology for PROMISE score development

Outcome measures

Two binary outcome measures were used; 1) Death before 3 months (91 days) from diagnosis for the survival analysis and 2) Pleurodesis success at 3 months, defined on objective and identical clinical criteria (see original publications for details)[7,8].

Candidate predictors

A total of 25 candidate predictors were assessed, including baseline demographics (age and sex), malignancy type, prior chemotherapy or radiotherapy, severity of illness, serum laboratory values, and pleural fluid proteins. Candidate predictors were selected based on their biological role.

Missing data

There were no missing data for the outcome measures. Among the predictors, data were missing ranging from 0 to 62% (supplementary appendix page 5). Missing data were imputed via multiple imputation using the chained equations and predictive mean matching (with 50 imputations). All available data was included in the imputation score.

Variable selection, model fitting, performance and discrimination

A two-stage process was used, first to select variables for the final model, and then to estimate the coefficients for the final model. Due to the large number of candidate predictors relative to the number of deaths models were developed with Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis using the GLMNET package in R (supplementary appendix page 3). Model discrimination was assessed using the c-statistic which quantifies the ability of model to differentiate between those who will die within 3 months, and those who will not. Model calibration was assessed through plots of the predicted versus observed mortality, and reporting of the calibration slope (which equals one for a perfectly calibrated

model). The Brier score was calculated as a general measure of predictive accuracy. Nagelkerke's R squared, a measure of overall performance was calculated. Each measure was calculated on each of the 50 imputed datasets after the second stage LASSO regression process. These performances were averaged to give the 'apparent' model performance.

Internal validation of PROMISE study model

Internal validation of the model development process was carried out using a bootstrap resampling process to provide an unbiased estimate of model performance. The original dataset was resampled with replacement to obtain a dataset of the same size. The full modelling procedure was applied, including multiple imputations. Further information about the statistical methodology can be found in the online supplementary appendix (pages 3-4).

Simplified model development

Using the methods outlined by Sullivan et al[17], a simplified version of the final model was created which was implementable in real life practice, whilst maintaining predictive performance and the ability to calculate absolute risk. All continuous predictors were categorised, and a reference value selected within each category, (usually the mid-point). Based on this, points were assigned to each category of each predictor. A simple table was produced to transform all possible total point scores into an absolute risk.

External validation and LENT score calculation

Using two independent and external datasets (the SIMPLE study[11] and Oxford Pleural Biobank), the final PROMISE score was applied and used to calculate an absolute risk for each subject. Based on this risk, performance measures were

calculated. The LENT score[7] was calculated in patients for direct comparison. The study was reported following the TRIPOD statement[18].

Statistics

Data summarised using mean or median and standard deviation or interquartile range. Differential protein expression was assessed with either t-test or Mann–Whitney U test (two-tailed $P<0{\cdot}05$). Specific statistical methods used on candidate predictors' selection or model development were described in previous sections. Analyses were done on the Statistical Package for the Social Sciences v24·0 (IBM, Armonk, NY) and Prism v7·0 (GraphPad, San Diego, USA).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit the manuscript for publication. The corresponding author had full access to all the data in the study and carries the final responsibility for the decision to submit for publication.

Ethical approval for the PROMISE study

Ethical and regulatory approval for the study was obtained by the South Central Oxford A Research Ethics Committee (REC reference number 15/SC/0186). Ethics approval for sample analysis for external validation was obtained before the laboratory investigations (Oxford Radcliffe Biobank ethics committee reference 17/A078).

RESULTS

Generation of a MPE proteomic data set from the targeted interventions for malignant pleural effusion (TIME-2) discovery dataset identifies candidate biomarkers of survival and pleurodesis success

The TIME-2 database was used for the discovery phase. Study participants were stratified based on survival and on pleurodesis outcome. For survival, patients were categorised as having poor (< 3 months) or good (≥ 3 months) survival based on median survival of the TIME-2 cohort (127 days) and on clinical decision for timing of providing best supportive care for MPE. The cut-off point of 3 months was selected based on clinical judgement, as there is paucity of data correlating days of survival that justify treatment with best supportive care instead of targeted treatment. Subjects were categorised to pleurodesis success or failure based on data at 3 months post pleural intervention. Demographic characteristics are presented in supplementary appendix (page 7). In total, we identified 1,250 proteins (95% minimum protein probability and minimum of 2 peptides) and 1,150 proteins for the survival and pleurodesis groups respectively (Figures 1A, 1B). Detailed information for the LC-MS/MS data is given in online supplement. 393 proteins and 93 proteins were statistically significant (t-test False Discovery Rate adjusted p value for multiple comparisons: q-value<0·05) differentially expressed with an absolute fold change greater than 2 between patients with poor and good survival and patients with pleurodesis success and failure respectively.

To corroborate the impact of the statistically significant differentially expressed proteins of the two datasets in an unbiased fashion, principal component analysis (PCA) was applied. In the survival dataset, the poor and good survival groups clearly

separated into two different clusters (Figure 1C). However, in the pleurodesis dataset, success and failure groups failed to separate and did not demonstrate distinct protein profiles (Figure 1D).

For the survival analysis, nine of 393 statistically significant differentially expressed proteins were identified whose biological role signalled a probable biomarker potency: gelsolin (GSN), metallopeptidase inhibitor 2, fibulin 3, laminin beta 1, extracellular matrix protein 1, versican (VCAN), macrophage migration inhibitory factor (MIF), galectin 1, and galectin 3. For the pleurodesis analysis, three out of 93 proteins were identified vascular cell adhesion protein 1 (VCAM1), matrix metalloproteinase 9 (MMP9), and angiopoietin-like 4 (ANGPTL4).

Pathway and an upstream regulator analysis using the IPA software, of the statistically significant proteins were performed. For the survival dataset, eight proteins were detected with probable therapeutic utility based on their biological role: tissue inhibitor metalloproteinase 1 (TIMP1), platelet-derived growth factor (PDGF), vascular endothelial growth factor (VEGF), cadherin-1 (CDH1), interleukin 4 (IL4), hypoxia-inducible factor 1 alpha (HIF1a), fibulin 3 and osteopontin (secreted phosphoprotein 1, SPP1). For the pleurodesis dataset, four proteins were identified as potential biomarkers: tumour necrosis factor alpha, (TNFα), tumour necrosis factor beta (TNFβ), interleukin 6 (IL6), and fibroblast growth factor 2 (FGF2).

Internal protein validation in the TIME-2 database of LC-MS/MS measured peptides

To validate LC-MS/MS results, protein levels of candidate biomarkers of all TIME-2 dataset samples were measured with either ELISA or Luminex assays

(supplementary appendix pages 7 and 8). For the survival analysis, we validated three of nine from the proteomic discovery experiment (GSN *P*=0.01, MIF *P*=0.03, and VCAN *P=0.02*). For the pleurodesis analysis, none of the candidate biomarkers could be validated (*P* > 0·05).

External protein validation in the TIME-1 and TIME-3 datasets identifies GSN, MIF, VCAN and TIMP1 as robust biomarkers of survival in MPE

To validate the biomarker utility of candidate proteins, expression was measured in two separate pleural fluid datasets (TIME-1 and TIME-3), which included 325 and 71 study patients respectively. Patients were categorized based on survival and pleurodesis outcomes and all biomarkers validated in the TIME-2 database (LC-MS/MS and IPA identified proteins) were measured in pleural fluid. All three proteins (GSN, MIF, and VCAN) from LC-MS/MS identified as prognostic biomarker candidates by LC-MS/MS showed a statistical significance in discriminating patients based on survival (Figure 2A-2D). A single upstream regulator protein from the IPA (TIMP1) showed differential expression that was statistically significant between the good and poor survival groups, whereas PDGF, VEGF, CDH1, IL4, HIF1a, FBLN 3, and SPP1 measurements did not differ significantly. All therapeutic (pleurodesis) biomarkers from the pathway analysis (TNFα, TNFβ, IL6, and FGF2) did not pass validation in the TIME-1 and TIME-3 datasets (*P* > 0·05).

Clinical and biological PROMISE model development and performance

We next sought to identify prognostic factors of survival in MPE. In the development dataset, 137/435 subjects died within 3 months (31%), and a total of 25 potential predictive factors including clinical, radiological (chest X ray findings) and biological

data (TIMP1, GSN, VCAN, MIF) were considered. For continuous risk factors, univariate analyses did not reveal any non-linear relationships with outcome. After LASSO regression, seven variables were retained in the final clinical score and eight in the biological score. After adjustment for optimism, the final clinical score achieved a C-statistic of 0·78, and included use of prior chemotherapy and radiotherapy, baseline performance status (PS), cancer type, haemoglobin, white cell count (WBC), C-reactive protein (CRP), and pleural fluid TIMP1 levels (Table 1). The biological score showed similar discrimination, with a C-statistic of 0·77, after adjusting for optimism. A simplified scoring system was derived for both the clinical and biological scores. (Table 2A and B). Both simplified scoring systems (clinical and biological) demonstrate good agreement with the full models (supplementary appendix pages 13 and 14).

Prospective assessment of PROMISE score in SIMPLE and Pleural Biobank databases

The external validation dataset comprised 162 subjects, of whom 58 died before the 3 month hallmark (36%) using SIMPLE study and Pleural Biobank databases. Discrimination was high and similar for full and simplified PROMISE models, at 0·90 (Table 1). Additionally, performance was similar for the full and simplified models. Both models were shown to be under-fitted, with calibration slopes greater than one (supplementary appendix pages 15 and 16). Based on the risk of three month death the PROMISE score can be classified to the following categories A: <25%, B: 25%-<50%, C: 50%-<75% and D: ≥75%. A nomogram for each score (clinical, biological) is shown in supplementary appendix (page 17)

Comparison of PROMISE with LENT score

The clinical and biological PROMISE scores and the LENT score were validated using complete case data, which included 192 subjects, of whom 62 died before 3 months (33%). The C-statistic to discriminate between those who do or do not die before 3 months was 0·75 (95% CI 0.68-0·81) for the LENT score (supplementary appendix page 12). As the LENT score was developed using Cox-regression and not a logistic model, Harrell's C-statistic was also calculated, and was 0·62 (95% CI 0·58-0·66). Kaplan-Meier estimates demonstrated good discrimination of survival for all three models and revealed that PROMISE scores performed better than LENT despite lower sample sizes (Figures 3B-C).

DISCUSSION

This is the first translational assessment combining discovery technology with high-quality clinical data from five different datasets for the development of a prognostic score for MPE. The results demonstrate the discovery, assessment and prospective validation of a novel prognostic score, including clinical data with pleural fluid biomarkers.

All parameters included in the PROMISE score are independently associated with survival, and thus the identified markers permit some speculation as to their biological role in survival in MPE. Those treated with previous chemotherapy and radiotherapy may have poorer prognosis due to the development of more aggressive cancer cell subpopulations that develop post-treatment. WBC and CRP are markers of systemic inflammatory responses, which are well associated with poor tumour-specific immunity[7,19]. Decreased haemoglobin is likely to indicate a more advanced stage of disease[20,21]. Performance status and type of primary cancer are well known factors, which are associated with prognosis in all cancers[7].

This study has demonstrated that pleural fluid TIMP1 levels are associated with survival, even when all of the above clinical factors are accounted for. TIMP1 is a glycoprotein which regulates the structure of the extracellular matrix, and previous studies have suggested that TIMP1 promotes cellular proliferation and anti-apoptotic activity[22-24]. Alongside with TIMP1, GSN, VCAN and MIF have been identified as potential biological factors correlated with MPE survival. These factors may suggest novel potential therapeutic targets in MPE with preclinical data to support it.

The development of the PROMISE score is a significant step forward in the precise management of MPE and could potentially aid patients and clinicians by enabling selection of optimal management strategy of the effusion. Evidence suggest that clinical judgement alone is imprecise at estimating patient survival, highlighted by the fact that physicians are ineffective in excluding study participants with poor prognosis from large clinical trials[8]. Taken into consideration that future confirmatory studies are required, the PROMISE score (either clinical or biological) has the potential to be used in everyday clinical practice as a method to improve patient management, improve associated healthcare costs and as an enrichment strategy for future clinical trials. Individuals with a PROMISE score category A (with less than 25% absolute risk of death) correlates with a particularly good prognosis, and these patients could be selected for more aggressive pleural and potential oncological or surgical management. In patients with a PROMISE score category D (at least 75% absolute risk of death), it is reasonable to offer minimally invasive procedures aimed at symptom control (e.g. therapeutic aspiration or indwelling pleural catheter insertion), best supportive care and a strategy to spend as few days as possible in hospital. Although clear recommendations cannot be given for patients with scores in categories B and C, the PROMISE study score provides a personalised absolute risk of death which will can be openly discussed during clinical consultation, and can be used as a basis for rational patient choices of further treatments offered.

Although the same experimental design to identify biomarkers of pleurodesis outcome was deployed, this was not successful. These results highlight the importance of the validation process during the biomarkers development and could be explained by the fact that a high-throughput technique was only applied on pleural fluid samples before pleurodesis. In the future, a comparison of pleural fluid before

and after pleurodesis from the same patient would potentially provide more information for a pathway analysis and identify biomarkers of pleurodesis success. Currently a multicentre, multinational clinical trial[11] is recruiting MPE patients that will provide samples before and after pleurodesis for further analysis.

In addition to the prognostic score, the PROMISE study provides important information about the underlying biology of MPE and identifies four novel candidates likely involved in MPE pathogenesis (supplementary appendix page 18). LC-MS/MS data identified four proteins (TIMP1, GSN, VCAN, MIF) of significant importance in designing novel treatment targets. In light of the results, significant differences in the molecular pathways between patients with poor and good survival were identified. Four molecular pathways and upstream regulators that are differentially expressed in patients who had exceptional prognosis for MPE were demonstrated. There are existing experimental and clinical data regarding the potential impact of the upstream regulators, SPP1, FBLN 3, HIF1, and PDGF in MPE[25-28]. The results of PROMISE study suggest that further investigation is required to examine the potential clinical impact of inhibition of these molecular pathways.

The strengths of this study include its design consisting of a three-step approach (discovery, validation, prospective assessment), a large sample size, and unique clinical resource with robust outcome data. Including all patients in the analysis and accounting for missing data using multiple imputations ensured all data were used and that potential missing associations were minimised. Prognostic scores were validated using two independent data sets, which add strength to the findings. Calculation of the PROMISE study score compared to current evidence (LENT score) provides an improved clinical prognostic score based on performance

(supplementary appendix page 12) and an advanced stratification score with personalised information for patients with MPE. Additionally, all variables that are included in the clinical PROMISE score are readily available in the clinical setting and measurement of TIMP-1 by the biochemistry department of a hospital can be achieved within 4 hours. In the real life situation when one parameter is missing, we would recommend using the last available measurement of the variable for the score. The important information that the PROMISE score provides compared to the LENT score justifies the extra variables that are required.

There are limitations to this study. The derived score is only applicable to patients with confirmed MPE in whom a pleural procedure is intended, and this precluded the general applicability on patients with “paramalignant” effusion or suspected MPE. Information for time from cancer diagnosis and treatment post score calculation were not included in the multivariate analysis, however, all patients were recruited at the time of definitive pleural intervention, and this can be considered an important single point clinically. A biological factor was included in the score which potentially limits the pragmatic application of the PROMISE score in every hospital practice. To mitigate this, a simplified score using only clinical information has been proposed, with reasonable performance.

To our knowledge this is the largest study in the literature with a systemic approach for identification of biomarkers and a prognostic score for MPE. The PROMISE study score is the first validated risk stratification system combining a discovered pleural fluid biomarker (TIMP1) with clinical information (previous chemotherapy or radiotherapy, haemoglobin, CRP, white cell blood count, performance status and type of primary cancer). It is therefore a robust, clinically relevant prognostic score

that can be applied immediately, and will provide important information on patient prognosis and guide the selection of appropriate management strategies.

CONTRIBUTORS

IP and NMR designed the study. MLT, PDC, RF and BMK performed the LC-MS/MS. IP, NIK, PDC, AS and HBS analysed the LC-MS/MS data. IP and NIK performed the protein measurement assays. IP, RB, RA, RJH, RM, NM, TD, MD, IDP, HIP, NM, NMR collected data. SG and GSC performed the statistical analysis and model development. IP, NIK and GTS analysed the data and created the figures. IP, NIK, RB, RR, AY, RF, AJE, RMM, JPC, RPH, TD, IDP, SP, HEGP, CC, NM, GTS, NMR drafted and reviewed the manuscript. All authors approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

ACKNOLEDGEMENTS

The authors would like to express their gratitude to Oxford Respiratory Trials Unit, University of Oxford for the support on PROMISE study and Dr Climent Casals-Pascual for guidance on biomarkers development. The authors thank Dr. Yanchun Peng and Dr. Xuan Yao from the Centre for Translational Immunology, University of Oxford for their support and guidance on Luminex assays. Prof Sam M Janes and Prof Maria G Belvisi for their guidance during the PROMISE study development. The authors thank as well the investigators of the TIME-1, TIME-2, TIME-3 and SIMPLE clinical trials. I Psallidas is the recipient of a REPSIRE2 European Respiratory Society Fellowship (RESPIRE2 – 2015–7160). NI Kanellakis has received a Short-

term Research Fellowship from the European Respiratory Society (STRTF 2015-9508). RJ Hallifax is funded by a Clinical Training Fellowship from the Medical Research Council (MR/L017091/1). GTS is funded by the European Research Council (grant numbers 260524 and 679345). NM Rahman is funded by the National Institute Health Research (NIHR) Oxford Biomedical Research Centre. We would like to thank as well our funding sources: Medical Research Funding, Slater & Gordon Research Fund and Oxfordshire Health Services Research Committee Research Grants.

REFERENCES

1. Roberts ME, Neville E, Berrisford RG, Antunes G, Ali NJ, Group BTSPDG. Management of a malignant pleural effusion: British Thoracic Society Pleural Disease Guideline 2010. *Thorax* 2010; **65 Suppl 2**: ii32-40.
2. Davies HE, Lee YC. Management of malignant pleural effusions: questions that need answers. *Current opinion in pulmonary medicine* 2013; **19**(4): 374-9.
3. Lee YC, Light RW. Management of malignant pleural effusions. *Respirology* 2004; **9**(2): 148-56.
4. Psallidas I, Kalomenidis I, Porcel JM, Robinson BW, Stathopoulos GT. Malignant pleural effusion: from bench to bedside. *European respiratory review : an official journal of the European Respiratory Society* 2016; **25**(140): 189-98.
5. Xia H, Wang XJ, Zhou Q, Shi HZ, Tong ZH. Efficacy and safety of talc pleurodesis for malignant pleural effusion: a meta-analysis. *PLoS One* 2014; **9**(1): e87060.
6. Clive AO, Bhatnagar R, Psallidas I, Maskell NA. Individualised management of malignant pleural effusion. *The Lancet Respiratory medicine* 2015; **3**(7): 505-6.
7. Clive AO, Kahan BC, Hooper CE, et al. Predicting survival in malignant pleural effusion: development and validation of the LENT prognostic score. *Thorax* 2014; **69**(12): 1098-104.
8. Davies HE, Mishra EK, Kahan BC, et al. Effect of an indwelling pleural catheter vs chest tube and talc pleurodesis for relieving dyspnea in patients with malignant pleural effusion: the TIME2 randomized controlled trial. *JAMA* 2012; **307**(22): 2383-9.
9. Rahman NM, Pepperell J, Rehal S, et al. Effect of Opioids vs NSAIDs and Larger vs Smaller Chest Tube Size on Pain Control and Pleurodesis Efficacy Among Patients With Malignant Pleural Effusion: The TIME1 Randomized Clinical Trial. *Jama* 2015; **314**(24): 2641-53.
10. Mishra EKBBMMD, Clive AO, Wills GH, et al. Randomised Controlled Trial of Urokinase versus Placebo for Non-draining Malignant Pleural Effusion. *American journal of respiratory and critical care medicine* 2017.
11. Psallidas I, Piotrowska HEG, Yousuf A, et al. Efficacy of sonographic and biological pleurodesis indicators of malignant pleural effusion (SIMPLE): protocol of a randomised controlled trial. *BMJ open respiratory research* 2017; **4**(1): e000225.
12. Fischer R, Kessler BM. Gel-aided sample preparation (GASP)--a simplified method for gel-assisted proteomic sample generation from protein extracts and intact cells. *Proteomics* 2015; **15**(7): 1224-9.
13. Team RC. A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria; 2015.
14. Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nature methods* 2015; **12**(2): 115-21.
15. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 2015; **43**(7): e47.
16. StroreyLab. R package to estimates q-values and false discovery rate quantities. 2011.
17. Sullivan LM, Massaro JM, D'Agostino RB, Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Statistics in medicine* 2004; **23**(10): 1631-60.
18. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). *Annals of internal medicine* 2015; **162**(10): 735-6.
19. Mei Z, Shi L, Wang B, et al. Prognostic role of pretreatment blood neutrophil-to-lymphocyte ratio in advanced cancer survivors: A systematic review and meta-analysis of 66 cohort studies. *Cancer treatment reviews* 2017; **58**: 1-13.
20. Allin KH, Nordestgaard BG. Elevated C-reactive protein in the diagnosis, prognosis, and cause of cancer. *Critical reviews in clinical laboratory sciences* 2011; **48**(4): 155-70.
21. Lu H, Ouyang W, Huang C. Inflammation, a key event in cancer development. *Molecular cancer research : MCR* 2006; **4**(4): 221-33.

22. Kim YS, Kim SH, Kang JG, Ko JH. Expression level and glycan dynamics determine the net effects of TIMP-1 on cancer progression. *BMB reports* 2012; **45**(11): 623-8.
23. Luparello C, Avanzato G, Carella C, Pucci-Minafra I. Tissue inhibitor of metalloprotease (TIMP)-1 and proliferative behaviour of clonal breast cancer cells. *Breast cancer research and treatment* 1999; **54**(3): 235-44.
24. Guedez L, Stetler-Stevenson WG, Wolff L, et al. In vitro suppression of programmed cell death of B cells by tissue inhibitor of metalloproteinases-1. *The Journal of clinical investigation* 1998; **102**(11): 2002-10.
25. Psallidas I, Stathopoulos GT, Maniatis NA, et al. Secreted phosphoprotein-1 directly provokes vascular leakage to foster malignant pleural effusion. *Oncogene* 2013; **32**(4): 528-35.
26. Elloul S, Vaksman O, Stavnes HT, Trope CG, Davidson B, Reich R. Mesenchymal-to-epithelial transition determinants as characteristics of ovarian carcinoma effusions. *Clinical & experimental metastasis* 2010; **27**(3): 161-72.
27. Langerak AW, De Laat PA, Van Der Linden-Van Beurden CA, et al. Expression of platelet-derived growth factor (PDGF) and PDGF receptors in human malignant mesothelioma in vitro and in vivo. *The Journal of pathology* 1996; **178**(2): 151-60.
28. Pass HI, Levin SM, Harbut MR, et al. Fibulin-3 as a blood and effusion biomarker for pleural mesothelioma. *The New England journal of medicine* 2012; **367**(15): 1417-27.

FIGURE LEGENDS

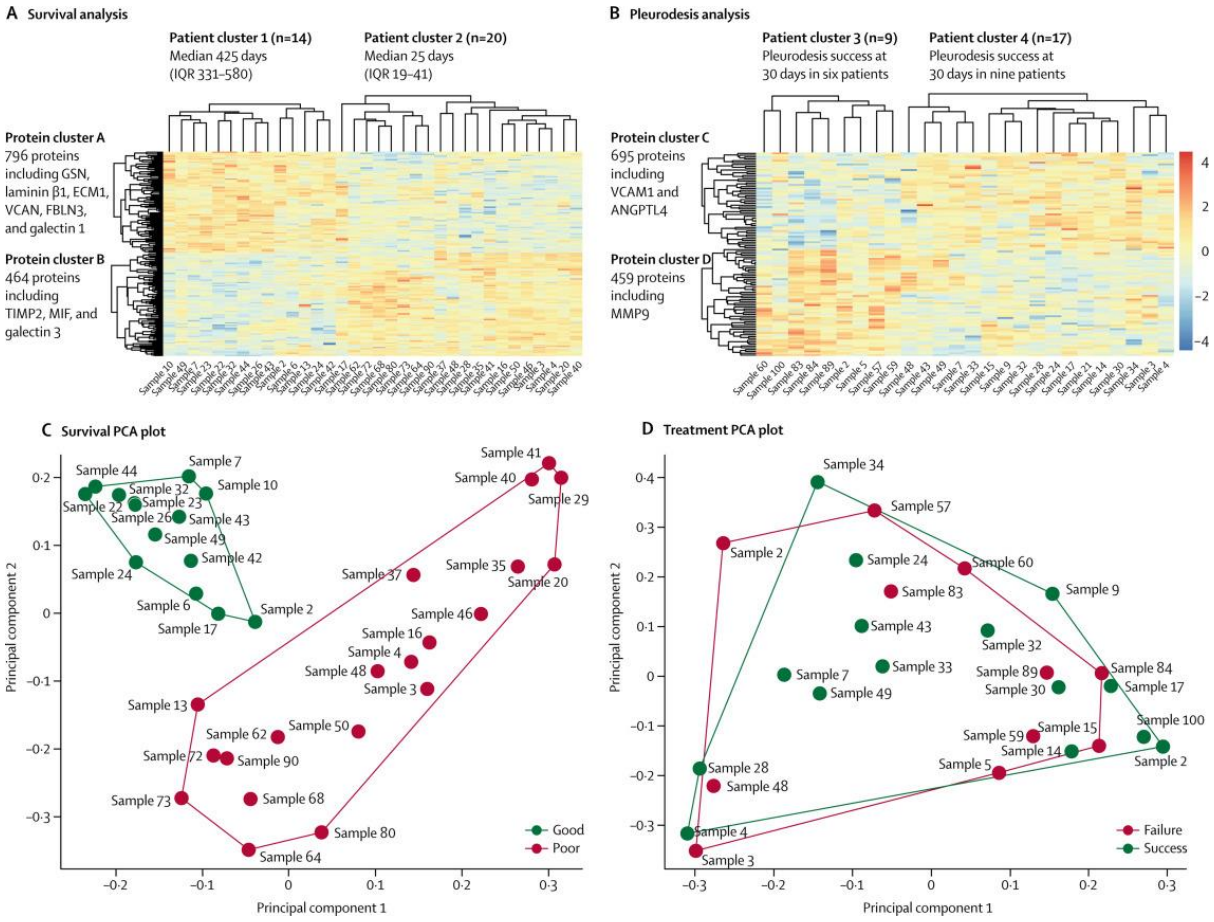

Figure 1. Proteomic analyses in the TIME-2 cohort. (A, B) Heatmaps of unsupervised hierarchical clustering for the survival (A) and pleurodesis (B) dataset samples. Red colour represents high expressive proteins and blue colour low expression proteins. **(C)** Principal component analysis (PCA) for the survival dataset. Red dots: Good Survival, Blue dots: Poor Survival. Interestingly, PCA separated the two groups into two different non-overlapping clusters. **(D)** PCA for the Pleurodesis dataset. Red dots: Failure, Blue dots: Success. Failure and Success groups failed to separate.

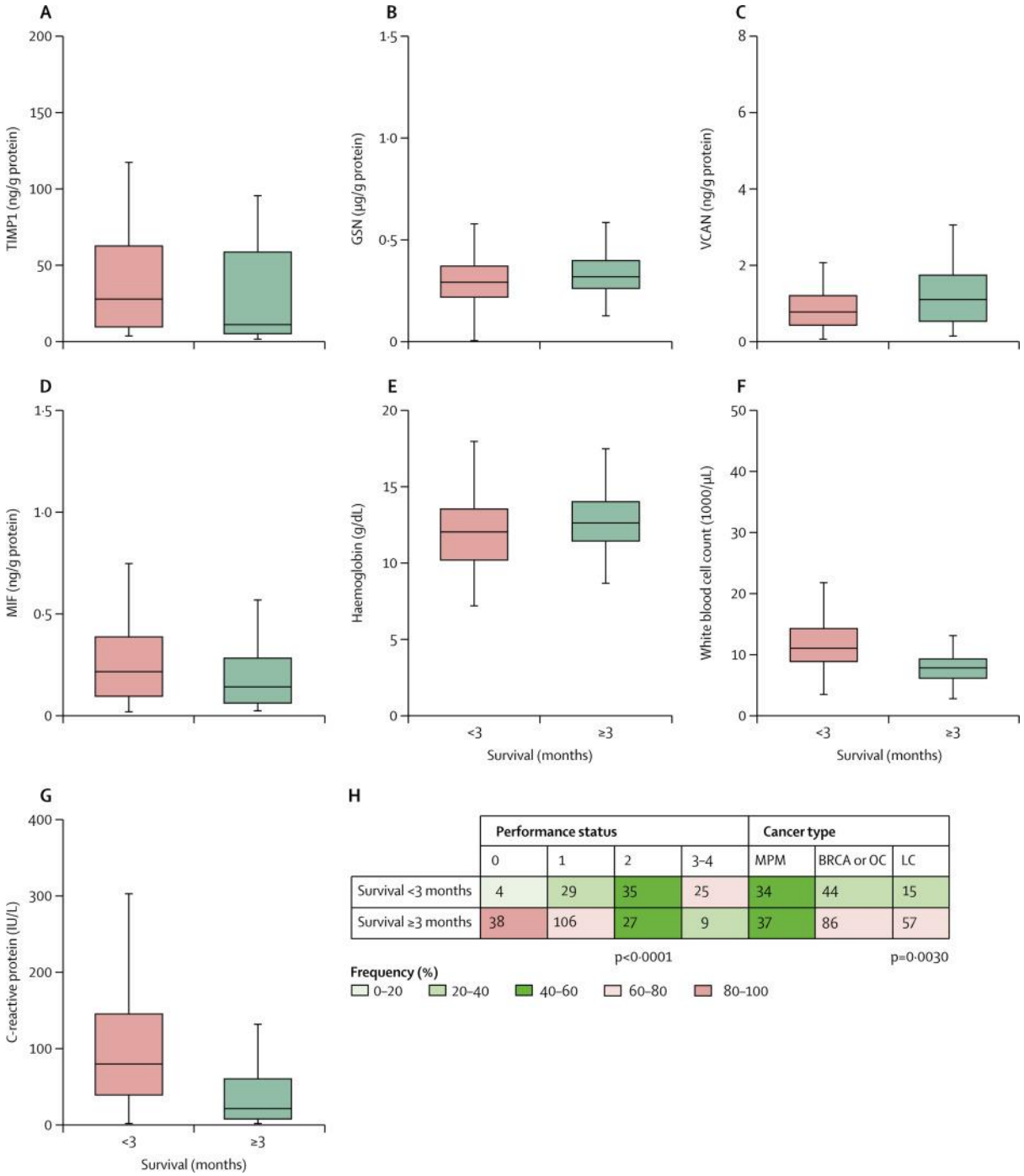

Figure 2. Validation of PROMISE components. (A-D) Relative protein expression (protein/total protein) of tissue inhibitor of metalloproteinase 1 (TIMP1, A), Gelsolin (GSN, B), Versican (VCAN, C), and macrophage inhibitory factor (MIF, D) in TIME-1, TIME-2 and TIME-3 datasets. **(E-H)** Selected clinical components of PROMISE in the TIME-1, TIME-2 and TIME-3 datasets. *P*, probability values for comparison

between long-term and poor survivors using t-test or Mann–Whitney U test (A-G) or chi square test (H).

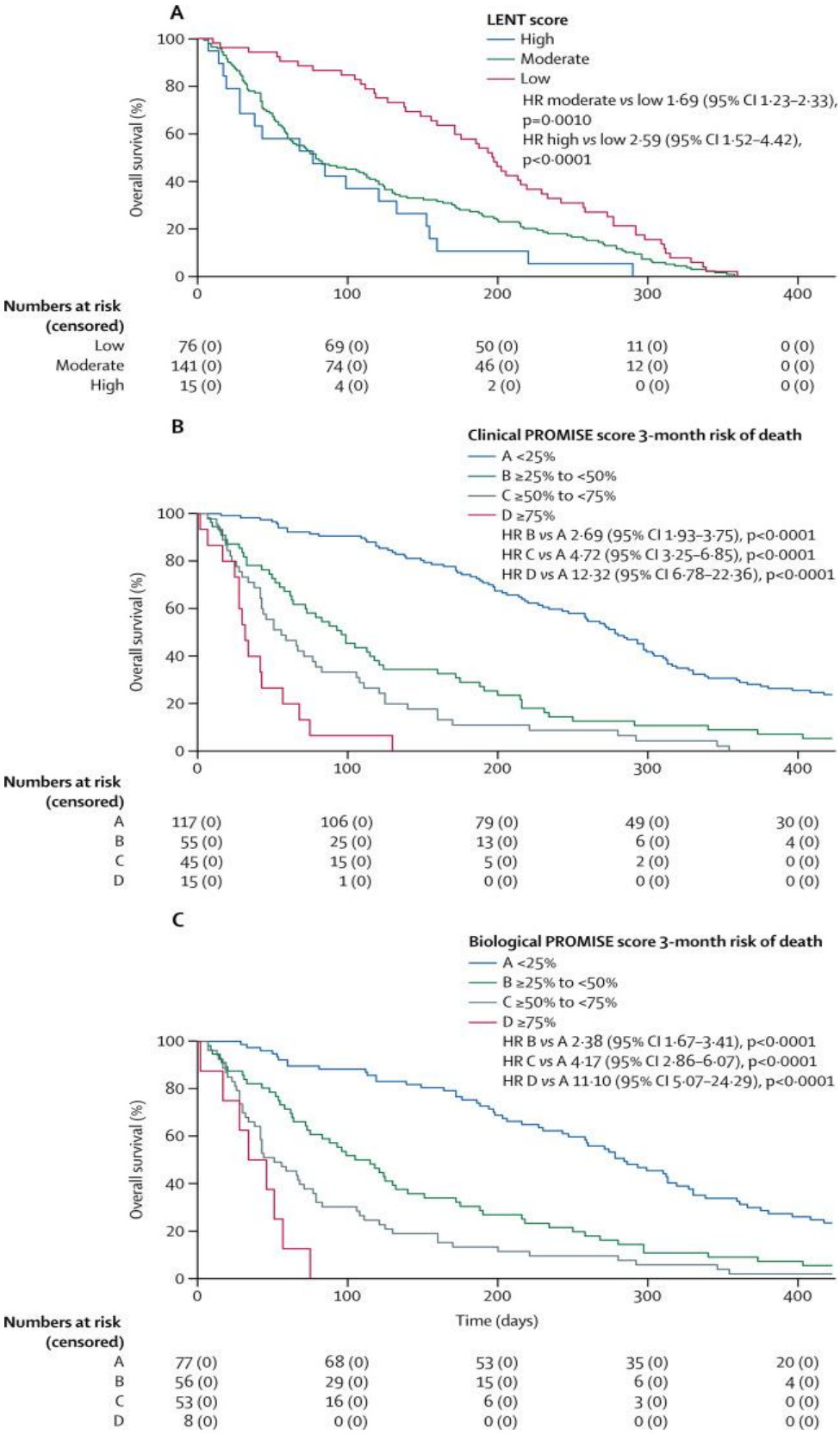

Figure 3. Performance of LENT score and PROMISE score in malignant pleural effusion patients. (A) Kaplan-Meier survival curves from patients with low, medium

and high LENT score. (B) Kaplan-Meier survival curves from patients classified into PROMISE score categories for the clinical and (C) biological PROMISE score. PROMISE score categories: A: <25% (blue line), B: 25%-<50% (green line), C: 50%-<75% (yellow line) and D: ≥75% (purple line).